

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 00:06:37 ; Search time 36 Seconds

(without alignments)  
2152.432 Million cell updates/sec

Title: US-09-758-017b-2

Perfect score: 1 MLEKLGICORCISNNKRLPG.....SKANGLIKSTGEPINRAL 301

Sequence: 1 MLEKLGICORCISNNKRLPG.....SKANGLIKSTGEPINRAL 301

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1045404 seqs, 257433775 residues

Total number of hits satisfying chosen parameters: 1045404

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1591	99.0	301	9 US-09-758-017A-2	Sequence 2, Appli
2	1380	85.9	301	9 US-09-758-017A-4	Sequence 4, Appli
3	1046	65.1	304	14 US-10-038-010-12	Sequence 12, Appli
4	1044.5	65.0	292	14 US-10-106-698-5713	Sequence 5713, Ap
5	693.5	43.2	225	15 US-10-369-493-236	Sequence 236, App
6	678.5	42.2	225	15 US-10-369-493-21107	Sequence 21107, A
7	677.5	42.2	229	15 US-10-369-493-23507	Sequence 23507, A
8	631.5	39.3	219	14 US-10-260-877-14	Sequence 14, Appli
9	631.5	39.3	247	15 US-10-369-493-524	Sequence 524, App
10	617	38.4	214	15 US-10-369-493-15993	Sequence 15993, A
11	617	38.4	237	15 US-10-369-493-15618	Sequence 15618, A
12	615	38.3	216	15 US-10-369-493-13886	Sequence 13886, A
13	600.5	37.4	220	15 US-10-369-493-10331	Sequence 10331, A
14	594	37.0	217	15 US-10-369-493-11585	Sequence 11585, A
15	594	37.0	220	15 US-10-369-493-14855	Sequence 14855, A

16	594	37.0	223	15 US-10-369-493-14189	Sequence 14189, A
17	588.5	36.6	225	15 US-10-369-493-16617	Sequence 16617, A
18	588	36.6	216	15 US-10-369-493-9179	Sequence 9179, Ap
19	587	36.5	256	15 US-10-369-493-17754	Sequence 17754, A
20	584.5	36.4	219	15 US-10-369-493-10737	Sequence 10737, A
21	575	35.8	212	15 US-10-369-493-15044	Sequence 15044, A
22	573.5	35.7	217	15 US-10-369-493-9434	Sequence 9434, Ap
23	571	35.5	223	15 US-10-369-493-17831	Sequence 17831, A
24	568.5	35.4	224	15 US-10-369-493-17465	Sequence 17465, A
25	566.5	35.3	225	15 US-10-369-493-23353	Sequence 23353, A
26	566	35.2	227	14 US-10-156-761-14539	Sequence 14539, A
27	554	34.5	359	15 US-10-369-493-1860	Sequence 1860, Ap
28	549	34.2	363	15 US-10-369-493-22663	Sequence 22663, A
29	537	33.4	329	15 US-10-369-493-7327	Sequence 7327, Ap
30	535.5	33.3	236	15 US-10-289-762-823	Sequence 823, App
31	514	32.0	198	15 US-10-369-493-4570	Sequence 4570, App
32	503.5	31.3	169	15 US-10-369-493-19125	Sequence 19125, A
33	502.5	31.3	225	15 US-10-369-493-8693	Sequence 8693, Ap
34	502.5	31.3	225	15 US-10-369-493-9579	Sequence 9579, Ap
35	489	30.4	309	15 US-10-369-493-3953	Sequence 3953, Ap
36	483.5	30.1	286	12 US-10-424-599-188911	Sequence 188911, A
37	454	28.3	384	15 US-10-369-493-13321	Sequence 13321, A
38	450.5	28.0	219	15 US-10-369-493-18312	Sequence 18312, A
39	380	23.6	233	10 US-09-882-227-322	Sequence 322, App
40	363.5	22.6	254	9 US-09-738-626-4961	Sequence 4961, App
41	359	22.3	223	15 US-10-369-493-8212	Sequence 8212, Ap
42	345	21.5	225	15 US-10-156-761-9066	Sequence 9066, Ap
43	275	17.1	72	9 US-09-925-301-1562	Sequence 1562, Ap
44	254.5	15.8	89	12 US-10-424-599-276019	Sequence 276019, A
45	158.5	9.9	154	12 US-10-424-599-188912	Sequence 188912, A

## ALIGNMENTS

```
RESULT 1
US-09-758-017A-2
Sequence 2, Application US/09758017A
Patent No. US2002015553A1
GENERAL INFORMATION:
APPLICANT: Lanes, Olav
APPLICANT: Willaen, Nils Peder
APPLICANT: Guddal, Per Henrik
APPLICANT: Gjellervik, Dag Rune
TITLE OF INVENTION: Cod uracil-DNA glycosylase, gene coding therefore,
TITLE OF INVENTION: recombinant DNA containing said gene or operative parts
TITLE OF INVENTION: thereof, a method for preparing said protein and the
FILE REFERENCE: U013209-3
CURRENT APPLICATION NUMBER: US/09758, 017A
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 2000 5428
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: 2000 0163
PRIOR FILING DATE: 2000-01-12
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 301
TYPE: PRT
ORGANISM: Gadus morhua
US-09-758-017A-2
Query Match 99.0%; Score 1591; DB 9; Length 301;
Best Local Similarity 99.3%; Pred. No. 4.5e-15;
Matches 299; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 MLEKLGICORCISNNKRLPGILITPOTLCFSKLMKITEKRLSSNVEQKTSPPQSLVQLE 60
Db 1 MLEKLGICORCISNNKRLPGILITPOTLCFSKLMKITEKRLSSNVEQKTSPPQSLVQLE 60
Cy 61 RMAKKAALDKITAKATPAAGFETWRRELAEPKRYFKOLMSFVADERSRHTVTPPAD 120
```

```

Db      61  RMAKKKALDKIRAKATPAGFGETWRELAEBEKEPYFKQLMSFVADERSHTVPPAD 120
Qy      121 QVTSWTEMCIDQVKVITLQDDPYHGNQAHGLCFVSQKVPVPPPSLVNIYKELCTDIDG 180
Db      121 QVTSWTEMCIDQVKVITLQDDPYHGNQAHGLCFVSQKVPVPPPSLVNIYKELCTDIDG 180
Qy      181 FKPHGHDLSGMAKQGVLLNNAVLTVAHQANSHKQSGMETFTDAVKMLSVNREGVPL 240
Db      181 FKPHGHDLSGMAKQGVLLNNAVLTVAHQANSHKQSGMETFTDAVKMLSVNREGVPL 240
Qy      241 LMGSYAHKKGATIDRKHNHVLQAVHPSLSAHRGFLGCKHFSKANGLLKLGSTEPINMRA 300
Db      241 LMGSYAHKKGATIDRKHNHVLQAVHPSLSAHRGFLGCKHFSKANGLLKLGSTEPINMRA 300
Qy      301 L 301
Db      301 L 301

```

## RESULT 2

```

US-09-758-017A-4
; Sequence 4, Application US/09758017A
; Patent No. US2002015573A1
; GENERAL INFORMATION:
; APPLICANT: Lanes, Olav
; APPLICANT: Willassen, Nils Peder
; APPLICANT: Gjeddesvik, Dag Rune
; TITLE OF INVENTION: Cod urecili-DNA glycosylase, gene coding therefore,
; TITLE OF INVENTION: recombinant DNA containing said gene or operative parts
; TITLE OF INVENTION: thereof, a method for preparing said protein and the
; FILE REFERENCE: U013209-3
; CURRENT APPLICATION NUMBER: US/09/758,017A
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 2000 5428
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 2000 0163
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Gadus morhua
US-09-758-017A-4

```

```

Query Match      85.9%; Score 1380; DB 9; Length 301;
Best Local Similarity 95.2%; Pred. No. 1.6e-138;
Matches 256; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      33 MITTPKLRSSNVEQKTSPPQSVQLEMAKKKALDKIRAKATPAGFGETWRELA 92
Db      33 VOITPKLRSSNVEQKTSPPQSVQLEMAKKKALDKIRAKATPAGFGETWRELA 92
Qy      93 BEKEPYFKQLMSFVADERSHTVPPADQVSWTEMCIDQVKVITLQDDPYHGNQAHG 152
Db      93 BEKEPYFKQLMSFVADERSHTVPPADQVSWTEMCIDQVKVITLQDDPYHGNQAHG 152
Qy      153 LCFVSQKVPVPPPSLVNIYKELCTDIDGFKPHGHDLSGMAKQGVLLNNAVLTVAHQAN 212
Db      153 LCFVSQKVPVPPPSLVNIYKELCTDIDGFKPHGHDLSGMAKQGVLLNNAVLTVAHQAN 212
Qy      213 SHKDGWETFTDAVKMLSVNREGVPLLMGSYAHKKGATIDRKHNHVLQAVHPSLSAH 272
Db      213 SHKDGWETFTDAVKMLSVNREGVPLFWGSYAHKKGATIDRKHNHVLQAVHPSLSAH 272
Qy      273 RGFLGCKHFSKANGLLKLGSTEPINMRA 301
Db      273 RGFLGCKHFSKANGLLKLGSTEPINMRA 301

```

## RESULT 3

```

US-10-038-010-12
; Sequence 12, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4/67A
; CURRENT APPLICATION NUMBER: US/10/038,010
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA glycosylase
; LOCATION: (1)..(304)
; OTHER INFORMATION:
US-10-038-010-12

```

```

Query Match      65.1%; Score 1046; DB 14; Length 304;
Best Local Similarity 67.7%; Pred. No. 7.6e-103;
Matches 195; Conservative 32; Mismatches 55; Indels 6; Gaps 3;

```

```

Qy      19 PG---LLIPQTLCSKLMKIRPKKLRSSNVQKT--SSPQSVQLEMAKKKALDKI 73
Db      18 PCKGFLQLSLQCDHQAATAKAARAGQEPGTPSP--LSAQDLRIQNRKAAALLRL 76
Qy      74 RAKATPAGFGETWRELAEBEKEPYFKQLMSFVADERSHTVPPADQVSWTEMCID 133
Db      77 AARVVPVGFSGSMKKHLSGEGKPYFTKMGFVAERKHVTVPPHQVFTWQMCID 136
Qy      134 VKVITLQDDPYHGNQAHGLCFVSQKVPVPPPSLVNIYKELCTDIDGFKPHGHDLSGMA 193
Db      137 VKVITLQDDPYHGNQAHGLCFVSQKVPVPPPSLVNIYKELCTDIDGFKPHGHDLSGMA 196
Qy      194 KQGVLLNNAVLTVAHQANSHKQSGMETFTDAVKMLSVNREGVPLLMGSYAHKKGAT 253
Db      197 KQGVLLNNAVLTVAHQANSHKQSGMETFTDAVSWLNQNSNGVFLLMGSYAHKKGAT 256
Qy      254 DRKHNVLQAVHPSLSAHRGFLGCKHFSKANGLLKLGSTEPINMRA 301
Db      257 DRKHNVLQAVHPSLSVVRGFFGCRHFSKTNELQKSGKKPIDWKL 304

```

```

RESULT 4
US-10-106-698-5713
; Sequence 5713, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: P0005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 5713
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5713

```

```
Query Match 65.0%; Score 1044.5; DB 14; Length 292;
Best Local Similarity 69.3%; Pred. No. 1e-102;
Matches 192; Conservative 31; Mismatches 51; Indels 3; Gaps 2;

Qy 27 LCFSLMKITPRLKLRSSNVEOKT--SSPOLSVBOLERMAKKKAAALDKIRAKATPAGFGE 84
Db 17 LCDHDLQAIIPAKKAPAGQBERGTPSSP-LSABOLDLQIKKAAALRLAARNVPVGFGE 75

Qy 85 TWRELAAEFEKPYFKOLMSFVADERSRHVYPADQVYSWTEKCDIQDKVYVILGDDP 144
Db 76 SWKHLSGEGFGKPYFKLMGFVAERKHYVYPPHGVFTWQCDIKDKVYVILGDDP 135

Qy 145 HGPNOAHGLCESVQKPPPPPSLVNIYKELCTDIDGKHPHGLSGMAKGVLLNNAV 204
Db 136 HGPNOAHGLCESVQKPPPPPSLVNIYKELCTDIEDVPHGHGDLSGMAKGVLLNNAV 195

Qy 205 TVRAHQANSHKDRKMETFTDAVIMKLSVNRGCVFLMGSAHKKGATIDKRRHVLQAV 264
Db 196 TVRAHQANSHKDRKMETFTDAVIMKLSVNRGCVFLMGSAHKKGATIDKRRHVLQAV 255

Qy 265 HPSPLSAHNGFLGCKHFSKANGLLKLSGTEPINRAL 301
Db 256 HPSPLSAHNGFLGCKHFSKANGLLKLSGTEPINRAL 292

RESULT 5
US-10-369-493-236
; Sequence 236; Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 236
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-236

Query Match 43.2%; Score 693.5; DB 15; Length 225;
Best Local Similarity 60.5%; Pred. No. 2.2e-65;
Matches 130; Conservative 29; Mismatches 55; Indels 1; Gaps 1;

Qy 85 TWRELAAEFEKPYFKOLMSFVADER-SRHVYPADQVYSWTEKCDIQDKVYVILGDDP 143
Db 6 TMHNVINENKQAPFTDTLTVANERKAGKITYPQDVFNAFRYTELADVKVILGDDP 65

Qy 144 YHGNQAHGLCFVSQKVPVPPPSLVNIYKELCTDIDGFKPHGHDLSGMAKGVLLNNAV 203
Db 66 YHGNQAHGLAFSQPQIPAPPSLVNMYKELESIDAGQRNHCGLSMKQGVLLNTV 125

Qy 204 LTVRAHQANSHKDRKMETFTDAVIMKLSVNRGCVFLMGSAHKKGATIDKRRHVLQAV 263
Db 126 LTVRGNVASHANLGWETFTDKVIAINEHRHGVIFLLMGSHAQKGFINTQHHVILKA 185

Qy 264 VHSPLSAHNGFLGCKHFSKANGLLKLSGTEPINR 298
Db 186 PHSPLSAHNGFLGCKHFSKANGLLKLSGTEPINR 220

RESULT 6
US-10-369-493-21107
; Sequence 21107; Application US/10369493
```

```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21107
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21107

Query Match 42.2%; Score 678.5; DB 15; Length 225;
Best Local Similarity 59.1%; Pred. No. 8.9e-64;
Matches 127; Conservative 31; Mismatches 56; Indels 1; Gaps 1;

Qy 85 TWRELAAEFEKPYFKOLMSFVADER-SRHVYPADQVYSWTEKCDIQDKVYVILGDDP 143
Db 6 TMHNVINENKQAPFTDTLTVANERKAGKITYPQDVFNAFRYTELADVKVILGDDP 65

Qy 144 YHGNQAHGLCFVSQKVPVPPPSLVNIYKELCTDIDGFKPHGHDLSGMAKGVLLNNAV 203
Db 66 YHGNQAHGLAFSQPQIPAPPSLVNMYKELESIDAGQRNHCGLSMKQGVLLNTV 125

Qy 204 LTVRAHQANSHKDRKMETFTDAVIMKLSVNRGCVFLMGSAHKKGATIDKRRHVLQAV 263
Db 126 LTVRGNVASHANLGWETFTDKVIAINEHRHGVIFLLMGSHAQKGFINTQHHVILKA 185

Qy 264 VHSPLSAHNGFLGCKHFSKANGLLKLSGTEPINR 298
Db 186 PHSPLSAHNGFLGCKHFSKANGLLKLSGTEPINR 220

RESULT 7
US-10-369-493-23507
; Sequence 23507; Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23507
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-23507

Query Match 42.2%; Score 677.5; DB 15; Length 229;
Best Local Similarity 60.9%; Pred. No. 1.2e-63;
Matches 131; Conservative 23; Mismatches 60; Indels 1; Gaps 1;

Qy 85 TWRELAAEFEKPYFKOLMSFVADER-SRHVYPADQVYSWTEKCDIQDKVYVILGDDP 143
Db 6 TMHNVINENKQAPFTDTLTVANERKAGKITYPQDVFNAFRYTELADVKVILGDDP 65
```

```

Qy      144 YHEPNQAHGCFEVSQKPVPPPSLVNIYKELCTDIDCFKHPGHDLSGMKQGVLLINAV 203
Db      66 YHOPGQAHGAFESVRGCIAPPSLILNNYKELNTIPGFTPNHNGYLESMARQGVLLINTV 125
Qy      204 LTVRAHQANSHKDRGWETFTDAVIKMLSVNREGVFLMLGMSYAHKKGATTIDRRHHVLA 263
Db      126 LTVRAQQAASHASISLGMETFTDKVLSILNQHREGVFLMLGMSHAQKGAITIDQRHHVLA 185
Qy      264 VHPSPLSAHRGFLGCKHESKANGLLKLSGTEPINW 298
Db      186 PHPSPLSAHRGFPGCNHFLVLANQMLEORGETPIDW 220

RESULT 8
US-10-260-877-14
; Sequence 14, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE OF INVENTION: 'ESSENTIAL GENES'
; FILE REFERENCE: 6565.US.P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 219
; TYPE: PRT
; ORGANISM: H. influenzae
US-10-260-877-14

Query Match      39.3%; Score 631.5; DB 14; Length 219;
Best Local Similarity 57.1%; Pred. No. 8,9e-59;
Matches 124; Conservative 27; Mismatches 65; Indels 1; Gaps 1;

Qy      84 ETWRRELAAEFEKPYFKQMLMSFVADER-SRHTVYPADQVYSWTMCMCIDQYKVVILGOD 142
Db      2 KMTVDVIGTEKAQPFQHTLQGVHLARASGKTIYPPQEDVFNAFKYTAIEDVKVILGOD 61
Qy      143 PYHGNGOAHGLCFYSQKRPVPPPSLVNIYKELCTDIDGFKHPGHGDLSCMAKQGVLLINA 202
Db      62 PYHGNGOAHGLAFYSKPEVALPPPSLILNYKELTODISGFQMPNSNGVLVYMAEQGVLLINT 121
Qy      203 VLTVAHQANSHKDRGWETFTDAVIKMLSVNREGVFLMLGMSYAHKKGATTIDRRKHNVLA 262
Db      122 VLTVRGAHSHSIANIGMERFTDKVLAVLNHHKELVFLMGSHAQKGAITIDRRHLVLT 181
Qy      263 AVHPSPLSAHRGFLGCKHESKANGLLKLSGTEPINW 299
Db      182 APHPSPLSAHRGFPGCRHFSKTSNLSHSGIKRPIDW 218

RESULT 9
US-10-369-493-524
; Sequence 524, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

```

```

US-10-369-493-524
# ORGANISM: Deinococcus radiodurans
# TYPE: Prt
# SEQ ID NO: 524
# LENGTH: 247
# PRIOR FILING DATE: 2002-02-21
# PRIOR APPLICATION NUMBER: US 60/360,039
# CURRENT FILING DATE: 2003-02-28
# CURRENT APPLICATION NUMBER: US/10/369,493

Query Match      39.3%; Score 631.5; DB 15; Length 247;
Best Local Similarity 56.1%; Pred. No. 1.1e-58;
Matches 124; Conservative 26; Mismatches 70; Indels 1; Gaps 1

Dy      79  PAGEGETRRRLAAEFKPYFKOLMSFVADERSHTVPPADQVYSWTEMCIDIQVKVI 138
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      20  PANLPEDMQEALRLPEFSAPFYHETLDFLRGRKKYTIYPAPADVFNMLARYTPFGVKLI 79

Dy      139 LGDPPYHGPNQAHGLCSVQKRPVPSPSLVNIYELCTDIDGFKNPGHGLSGMAKQVYL 198
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      80  LGDPPYHGPNQAHGLCSFSVRGVAVPSPSLRIYKELTFEDIPGVAAPKKHGYLRSAHQGVYL 139

Dy      199 LLNVLVYRAQANSHSDRGWETTTDAVIKLSVNRGCVYVLLMGSAHKKGATITDRKH 258
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      140  LLNVLVYRAQANSHQSGKQWHEHTTDAVIKAVNKKERVAVFIILMGSAARKKKLITGRKH 199

Dy      259 HVLQAVHSPSLSAHRGFGCGHGFESKANGLLKSGTEPINR 299
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      200  VIESGHPSPLS-QYFFGTRPFTKIDALEKAGRGVEWO 239

```

```

RESULT 10
US-10-369-493-15993
; Sequence 15993, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15993
; LENGTH: 214
; TYPE: PRF
; ORGANISM: Xanthomonas campestris
US-10-369-493-15993

Query Match          38.4%; Score 617; DB 15; Length 214;
Best Local Similarity 53.5%; Pred. No. 3,1e-57;
Matches 115; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

QY      85 TWRRRLAAEFKPYRKQMLSPVADERSNHT-VYPPADGVSYMTWENCIDIQVKKVILGGDP 143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       1 SWKARVGEWILPQMCQELSAFLRQRKANARVFPPGPQIFAFAFDATPEQYKVVVLGGDP 60

QY      144 YHGEPNDAHGLCSYQKRPVPSPSLVNIYKELCTDIDGRKHGHGDLSCMAQGVILLNAV 203
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       61 YHGGGGAHGLCSYVPGVPVPPSLNIIYKEIQDDL-GIPRDHGLVMWARGVILLNAV 119

QY      204 LTVRAHQANSHDKRGMEFTDAVIVKILSVNREGVFTLMSGVAAHKGCATTIRKKEHHVLOA 263
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       120 LTVGGGAGAHQNKMBEGTTHVIVETLNSREGVFTLMSGVAAQSKGVITQAHHRVYKA 179

QY      264 VHPSPLSAHRGFLGCKHPSKANGILIKSGTEPINW 298

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2004, 00:06:37 ; Search time 36 Seconds

(without alignment)  
2152.432 Million cell updates/sec

Title: US-09-758-017b-4

Perfect score: 1 MIGOQHINSFSPVSKRRVS.....SKANGLIKSGTEPIINRAL 301

Sequence:

Scoring table:

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database:

Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1550	96.6	301	9 US-09-758-017A-4	Sequence 4, Appli
2	1421	88.5	301	9 US-09-758-017A-2	Sequence 2, Appli
3	1036	64.5	292	14 US-10-106-698-5713	Sequence 12, Appl
4	1036	64.5	304	14 US-10-038-010-12	Sequence 12, Appl
5	693.5	42.3	225	15 US-10-369-493-236	Sequence 216, App
6	678.5	42.2	229	15 US-10-369-493-21107	Sequence 21507, A
7	677.5	42.2	229	15 US-10-369-493-23507	Sequence 21507, A
8	631.5	39.3	219	14 US-10-260-877-14	Sequence 14, Appl
9	631.5	39.3	247	15 US-10-369-493-524	Sequence 524, App
10	617	38.4	214	15 US-10-369-493-15993	Sequence 15993, A
11	617	38.4	237	15 US-10-369-493-15618	Sequence 15618, A
12	615	38.3	216	15 US-10-369-493-13886	Sequence 13886, A
13	600.5	37.4	220	15 US-10-369-493-10331	Sequence 10331, A
14	594	37.0	217	15 US-10-369-493-11585	Sequence 11585, A
15	594	37.0	220	15 US-10-369-493-14855	Sequence 14855, A

16	594	37.0	223	15 US-10-369-493-14189	Sequence 14189, A
17	588.5	36.7	225	15 US-10-369-493-16617	Sequence 16617, A
18	588	36.6	216	15 US-10-369-493-9179	Sequence 9179, Ap
19	587	36.6	256	15 US-10-369-493-17754	Sequence 17754, A
20	584.5	36.4	219	15 US-10-369-493-10737	Sequence 10737, A
21	575	35.8	212	15 US-10-369-493-15044	Sequence 15044, A
22	573.5	35.7	217	15 US-10-369-493-9434	Sequence 9434, Ap
23	571	35.6	223	15 US-10-369-493-17831	Sequence 17831, A
24	568.5	35.4	224	15 US-10-369-493-17465	Sequence 17465, A
25	566.5	35.3	225	15 US-10-369-493-23353	Sequence 23353, A
26	566	35.3	227	14 US-10-156-761-14539	Sequence 14539, A
27	558	34.8	359	15 US-10-369-493-1860	Sequence 1860, Ap
28	556.5	34.7	363	15 US-10-369-493-22663	Sequence 22663, A
29	535.5	33.4	236	15 US-10-289-762-823	Sequence 823, App
30	528.5	32.9	328	15 US-10-369-493-7327	Sequence 7327, Ap
31	514	32.0	198	15 US-10-369-493-4570	Sequence 4570, Ap
32	503.5	31.4	169	15 US-10-369-493-19125	Sequence 19125, A
33	502.5	31.3	225	15 US-10-369-493-8693	Sequence 8693, A
34	502.5	31.3	225	15 US-10-369-493-9579	Sequence 9579, Ap
35	493	30.7	286	12 US-10-424-599-188911	Sequence 188911, A
36	488.5	30.4	309	15 US-10-369-493-3953	Sequence 3953, Ap
37	459	28.6	384	15 US-10-369-493-13321	Sequence 13321, A
38	450.5	28.1	219	15 US-10-369-493-18312	Sequence 18312, A
39	380	23.7	233	10 US-09-882-227-322	Sequence 322, App
40	363.5	22.6	254	9 US-09-738-626-4961	Sequence 4961, Ap
41	359	22.4	223	15 US-10-369-493-8212	Sequence 8212, Ap
42	345	21.5	225	14 US-10-156-761-9066	Sequence 9066, Ap
43	275	17.1	72	9 US-09-925-301-1562	Sequence 1562, Ap
44	254.5	15.9	89	12 US-10-424-599-276019	Sequence 276019, A
45	164	10.2	154	12 US-10-424-599-188912	Sequence 188912, A

## ALIGNMENTS

RESULT 1  
US-09-758-017A-4  
Sequence 4, Application US/09758017A  
Patent No. US2002015573A1  
GENERAL INFORMATION:  
APPLICANT: Millaen, Nils Peder  
APPLICANT: Millaen, Olav  
APPLICANT: Guddal, Per Henrik  
APPLICANT: Gjelleevik, Dag Rune  
TITLE OF INVENTION: Cod uracil-DNA glycosylase, gene coding therefore,  
TITLE OF INVENTION: recombinant DNA containing said gene or operative parts  
TITLE OF INVENTION: thereof, a method for preparing said protein and the  
TITLE OF INVENTION: use of said protein or said operative pa  
FILE REFERENCE: U013209-3  
CURRENT APPLICATION NUMBER: US/09/758, 017A  
CURRENT FILING DATE: 2001-01-10  
PRIOR APPLICATION NUMBER: 2000 5428  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: 2000 0163  
PRIOR FILING DATE: 2000-01-12  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 301  
TYPE: PRT  
ORGANISM: Gadus morhua  
US-09-758-017A-4  
Query Match 96.6%; Score 1550; DB 9; Length 301;  
Best Local Similarity 96.3%; Pred. No. 4.5e-147;  
Matches 290; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MIGOQHINSFSPVSKRRVSKELKTEKHAEEVOITPKKLSSNVEQTSPPQLSVEBLE 60  
DB 1 MIGOQHINSFSPVSKRRVSKELKTEKHAEEVOITPKKLSSNVEQTSPPQLSVEBLE 60  
QY 61 RMANKKAAADIKTAKATPAGFETWRELAEEKPEFKQLMSFVADERSRHTVYPPAD 120

```

Db      61  RMAKAKKALDKIRAKATPAGFGEFTWRRELAEPKPYFKQLMSFVADERSRRTVPPAD 120
QY      121  QVYSWTEMCDIQVKVVIIGQDPYHGNQAHGLCFSVQKVPVPPPSLVNIYKELCTDIDG 180
Db      121  QVYSWTEMCDIQVKVVIIGQDPYHGNQAHGLCFSVQKVPVPPPSLVNIYKELCTDIDG 180
QY      181  FKHPGHGDLGSMAGKQGVLLNNAVLTVRAHQANSHKRGWTFPDVATIKMLSVNREGVFL 240
Db      181  FKHPGHGDLGSMAGKQGVLLNNAVLTVRAHQANSHKRGWTFPDVATIKMLSVNREGVFL 240
QY      241  LMGSAHKKGATIDRRKHVLAQVHPSPLSAHGFGLCKHFSKANGLLKLGSTEPINMRA 300
Db      241  FMGSAHKKGATIDRRKHVLAQVHPSPLSAHGFGLCKHFSKANGLLKLGSTEPINMRA 300
QY      301  L 301
Db      301  L 301

```

## RESULT 2

```

US-09-758-017A-2
; Sequence 2, Application US/09758017A
; Patent No. US20020155573A1
; GENERAL INFORMATION:
; APPLICANT: Lanes, Olav
; APPLICANT: Willasen, Nils Peder
; APPLICANT: Guddal, Per Henrik
; APPLICANT: Gjelleevik, Dag Rune
; TITLE OF INVENTION: Cod urecili-DNA glycosylase, gene coding therefore,
; TITLE OF INVENTION: recombinant DNA containing said gene or operative parts
; TITLE OF INVENTION: thereof, a method for preparing said protein and the
; FILE REFERENCE: U013209-3
; CURRENT APPLICATION NUMBER: US/09/758,017A
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 2000 5428
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 2000 0163
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Gadus mornua
US-09-758-017A-2

```

```

Query Match      88.5%; Score 1421; DB 9; Length 301;
Best Local Similarity 98.9%; Pred. No. 4,2e-134;
Matches 266; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      33  VOITPKLRSSVNEQKTSPPQSLSVBQLEMAKKAALDKIRAKATPAGFGEFTWRRELA 92
Db      33  KMITPKLRSSVNEQKTSPPQSLSVBQLEMAKKAALDKIRAKATPAGFGEFTWRRELA 92
QY      93  EEKPEPKQMSFVADERSRHTVVPADQVYSWTEMCDIQVKVVIIGQDPYHGNQAHG 152
Db      93  EEKPEPKQMSFVADERSRHTVVPADQVYSWTEMCDIQVKVVIIGQDPYHGNQAHG 152
QY      153  LCFSVQKVPVPPPSLVNIYKELCTDIDGFKHPGHGDLGSMAGQGVLLNNAVLTVRAHQAN 212
Db      153  LCFSVQKVPVPPPSLVNIYKELCTDIDGFKHPGHGDLGSMAGQGVLLNNAVLTVRAHQAN 212
QY      213  SHKDRGWEFTPDVATIKMLSVNREGVFLMGSAHKKGATIDRRKHVLAQVHPSPLSAH 272
Db      213  SHKDRGWEFTPDVATIKMLSVNREGVFLMGSAHKKGATIDRRKHVLAQVHPSPLSAH 272
QY      273  RGFLGCKHFSKANGLLKLGSTEPINMRA 301
Db      273  RGFLGCKHFSKANGLLKLGSTEPINMRA 301

```

RESULT 3

```

US-10-106-698-5713
; Sequence 5713, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA00591
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5713
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5713

```

```

Query Match      64.5%; Score 1036; DB 14; Length 292;
Best Local Similarity 68.1%; Pred. No. 2e-95;
Matches 186; Conservative 35; Mismatches 50; Indels 2; Gaps 1;

```

```

QY      31  EKVQITPKLRSSVNEQKTSPPQSLSVBQLEMAKKAALDKIRAKATPAGFGEFTWR 88
Db      20  DHLQAIIPAKKAPAGEEBGTGTPSSPLSAEQLDRIQRNKAALLRLAANVPVGGESWKK 79
QY      89  ELAAFEKPYKQLMSFVADERSRHTVVPADQVYSWTEMCDIQVKVVIIGQDPYHGN 148
Db      80  HLSGFPGKPYFKLMGFABERKHYTVPPHGVFTWQMDIKDVKVIIGQDPYHGN 139
QY      149  QAHGLCFSVQKVPVPPPSLVNIYKELCTDIDGFKHPGHGDLGSMAGQGVLLNNAVLTVRA 208
Db      140  QAHGLCFSVQKVPVPPPSLVNIYKELCTDIDGFKHPGHGDLGSMAGQGVLLNNAVLTVRA 199
QY      209  HQANSHKDRGWEFTPDVATIKMLSVNREGVFLMGSAHKKGATIDRRKHVLAQVHPS 268
Db      200  HQANSHKDRGWEFTPDVATIKMLSVNREGVFLMGSAHKKGATIDRRKHVLAQVHPS 259
QY      269  LSAHGFGLCKHFSKANGLLKLGSTEPINMRA 301
Db      260  LSVYRGFGCRHFSKTNELLQKSGKPIDWKEL 292

```

## RESULT 4

```

US-10-038-010-12
; Sequence 12, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B467A
; CURRENT APPLICATION NUMBER: US/10/038,010
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA glycosylase
; LOCATION: (1)..(304)
; OTHER INFORMATION:
US-10-038-010-12

```



```

Query Match      64.5%; Score 1036; DB 14; Length 304;
Best Local Similarity 68.1%; Pred. No. 2.1e-95;
Matches 186; Conservative 35; Mismatches 50; Indels 2; Gaps 1;

Oy 31 EEVOITPKLRSSNVEQKTSPO--LSVEOLERMAKKKAAALDKIRAKATFAGGFTWR 88
D 32 DHQAIIPAKAQAQGEPTGTPSSPLSABQDRIQRKAAALLRLAARNTFVGSGSMK 91
Oy 89 ELAAEFKPYFKOLMSFVADERSHTYVPADQYVSWTEMCIDQVKKVILGDDPYHGN 148
D 92 HLGSEFGKPYFKILMGVAERKKHYTYVPRPHQVFTQWCDIKDVKKVILGDDPYHGN 151
Oy 149 QAHGLCFSVQKRPVPPPSLVNIYKELCTDIDGFRHGHGDLGSMAGVILLNAVLTVA 208
D 152 QAHGLCFSVQKRPVPPPSLVNIYKELCTDIEDFVPHGHGDLGSMAGVILLNAVLTVA 211
Oy 209 HQANSHDRGMEFTDAVIMKLSVNRGVFLLMGSAHKKGATIDRKHHVILQAVHASP 268
D 212 HQANSHKRGMEQTDVAVSNLQNSGLVFLMGSAHKKGATIDRKHHVILQAVHASP 271
Oy 269 LSAHRGFLGCKHFSKANGLLKLGSTEPINRAL 301
D 272 LSVYRGFGCRHFSKTNELLQSGKKCIDWKEL 304

RESULT 5
US-10-369-493-236
; Sequence 236, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 236
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-236

Query Match      43.2%; Score 693.5; DB 15; Length 225;
Best Local Similarity 60.5%; Pred. No. 3.7e-61;
Matches 130; Conservative 29; Mismatches 55; Indels 1; Gaps 1;

Oy 85 TWRELAAEFEKPYFKOLMSFVADER-SRHTVYPADQYVSWTEMCIDQVKKVILGDDP 143
D 6 TMDHVINERKQAPFYFTDLTYVANERKAGKITYPQDVFNAFRYTELADKKVILGDDP 65
Oy 144 YHGNQAHGLCFSVQKRPVPPPSLVNIYKELCTDIDGFRHGHGDLGSMAGVILLNAV 203
D 66 YHGNQAHGLAFSVQPGIPAPPSLVNMYKELSDIAGFORRNHGLGSMAGVILLNAV 125
Oy 204 LTVRAHQANSHKRGMEFTDAVIMKLSVNRGVFLLMGSAHKKGATIDRKHHVILQ 263
D 126 LTVRGNVASHANLGMETFTDKVIAINERHSGVIFLLMGSAHKKGATIDRKHHVILQ 185
Oy 264 VHPSPLSAHRGFLGCKHFSKANGLLKLGSTEPINW 298
D 186 PHPSPLSAHRGFLGCKHFSQANHLLEAGLASIDW 220

RESULT 6
US-10-369-493-21107
; Sequence 21107, Application US/10369493

```

```

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21107
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21107

Query Match      42.3%; Score 678.5; DB 15; Length 225;
Best Local Similarity 59.1%; Pred. No. 1.2e-59;
Matches 127; Conservative 31; Mismatches 56; Indels 1; Gaps 1;

Oy 85 TWRELAAEFEKPYFKOLMSFVADER-SRHTVYPADQYVSWTEMCIDQVKKVILGDDP 143
D 6 TMDHVINERKQAPFYFTDLTYVANERKAGKITYPQDVFNAFRYTELADKKVILGDDP 65
Oy 144 YHGNQAHGLCFSVQKRPVPPPSLVNIYKELCTDIDGFRHGHGDLGSMAGVILLNAV 203
D 66 YHGNQAHGLAFSVQPGIPAPPSLVNMYKELSDITGFSRPNHGLGSMAGVILLNAV 125
Oy 204 LTVRAHQANSHKRGMEFTDAVIMKLSVNRGVFLLMGSAHKKGATIDRKHHVILQ 263
D 126 LTVRGNVASHANLGMETFTDKVIAINERHSGVIFLLMGSAHKKGATIDRKHHVILQ 185
Oy 264 VHPSPLSAHRGFLGCKHFSKANGLLKLGSTEPINW 298
D 186 PHPSPLSAHRGFLGCKHFSQANHLLEAGLTPIDW 220

RESULT 7
US-10-369-493-23507
; Sequence 23507, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23507
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-23507

Query Match      42.2%; Score 677.5; DB 15; Length 229;
Best Local Similarity 60.9%; Pred. No. 1.5e-57;
Matches 131; Conservative 23; Mismatches 60; Indels 1; Gaps 1;

Oy 85 TWRELAAEFEKPYFKOLMSFVADER-SRHTVYPADQYVSWTEMCIDQVKKVILGDDP 143
D 6 TMDHVLAEKQAPFYFTDLTVASERQSVTIYPQDVFNAFRYTELADKKVILGDDP 65

```

```

QY 144 YHGNQAHGLCFSVQKVPVPPPSLVNIYKEICTDIDGFKHGGDLGMAKQGVLLNAV 203
DB 66 YHGGQAHGLAFSVRPGVPPPSLVNIYKELENTIPGFTPNHGYLSMARQGVLLNTV 125
QY 204 LTVRAHQANSHDKRGWETTDVAIKWLSVNRGCVFLLMGSAHKKGATIDRKHVLA 263
DB 126 LTVRAQAHGASHASLGWETFDKVISLNOHREGVVFLLMGSHAKKATIDKORHVLKA 185
QY 264 VHPSPLSAHRGFLGCKHFSKANGLLKLGSTEPINW 298
DB 186 PHPSPLSAHRGFGCNHFVLANQWLEQRGTEPIDM 220

```

## RESULT 8

```

US-10-260-877-14
; Sequence 14, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Heesler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE REFERENCE: 6565 US P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 219
; TYPE: PRT
; ORGANISM: H. influenzae
US-10-260-877-14

```

Query Match 39.3%; Score 631.5; DB 14; Length 219;  
Best Local Similarity 57.1%; Pred. No. 6e-55;

Matches 124; Conservative 27; Mismatches 65; Indels 1; Gaps 1;

```

QY 84 EWRRELAAEFEKPYFKOLMSFYADERSRHTVPPADQVYMTWEMCDIDQVYVILGOD 142
DB 2 KMTVDVIGTEKAPRYQHTLQOVHLARASGKTIYPPQEVNPAFKTAEDVKVILGOD 61
QY 143 YHGNQAHGLCFSVQKVPVPPPSLVNIYKEICTDIDGFKHGGDLGMAKQGVLLNAV 202
DB 62 PYHGPQAHGLAFSVRPGVPPPSLVNIYKEICTDISGFQMSNGVLYMAAQVLLNT 121
QY 203 VLTVRANQANSHDKRGWETTDVAIKWLSVNRGCVFLLMGSAHKKGATIDRKHVLA 262
DB 122 VLTVERGMAHSHANLGWETFDKVISLNOHREGVVFLLMGSHAKKATIDKORHVL 181
QY 263 AVHPSPLSAHRGFLGCKHFSKANGLLKLGSTEPINW 299
DB 182 PHPSPLSAHRGFGCNHFVLANQWLEQRGTEPIDM 220

```

## RESULT 9

```

US-10-369-493-524
; Sequence 524, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

```

```

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 524
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Deinococcus radiodurans
US-10-369-493-524

```

Query Match 39.3%; Score 611.5; DB 15; Length 247;  
Best Local Similarity 56.1%; Pred. No. 7.2e-55;

Matches 124; Conservative 26; Mismatches 70; Indels 1; Gaps 1;

```

QY 79 PAGFETWRRELAAEFEKPYFKOLMSFYADERSRHTVPPADQVYMTWEMCDIDQVYVIL 138
DB 20 PANLEPDMQELALPEFSAPYHELTDFLRQREKTYTTPPADVFNALRYTPLEGVKLI 79
QY 139 LGQDPYHGNQAHGLCFSVQKVPVPPPSLVNIYKEICTDIDGFKHGGDLGMAKQGV 198
DB 80 LGQDPYHGNQAHGLCFSVRPGVPPPSLVNIYKELENTIPGFTPNHGYLSMARQGV 139
QY 199 LTVRAHQANSHDKRGWETTDVAIKWLSVNRGCVFLLMGSAHKKGATIDRKHVLA 258
DB 140 LTVRAQAHGASHASLGWETFDKVISLNOHREGVVFLLMGSHAKKATIDKORHVL 199
QY 259 VLTVRANQANSHDKRGWETTDVAIKWLSVNRGCVFLLMGSAHKKGATIDRKHVLA 299
DB 200 VVIESGHSPLS-RQYFGTRPFPSKTNLEAKRGVEMQ 239

```

## RESULT 10

```

US-10-369-493-15993
; Sequence 15993, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15993
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15993

```

Query Match 38.4%; Score 617; DB 15; Length 214;  
Best Local Similarity 53.5%; Pred. No. 1.7e-53;

Matches 115; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

```

QY 85 TWRELAAEFEKPYFKOLMSFYADERSRHT-VYPPADQVYMTWEMCDIDQVYVILGOD 143
DB 1 SWKARVGEMLDQOMELSAFLRQKAVANAVPPROPFLAAMPDTPPEQVYVVLGOD 60
QY 144 YHGNQAHGLCFSVQKVPVPPPSLVNIYKEICTDIDGFKHGGDLGMAKQGVLLNAV 203
DB 61 YHGGQAHGLAFSVRPGVPPPSLVNIYKELENTIPGFTPNHGYLSMARQGVLLNAV 119
QY 204 LTVRAHQANSHDKRGWETTDVAIKWLSVNRGCVFLLMGSAHKKGATIDRKHVLA 263
DB 120 LTVRAQAHGASHASLGWETFDKVISLNOHREGVVFLLMGSHAKKATIDKORHVL 179
QY 264 VHPSPLSAHRGFLGCKHFSKANGLLKLGSTEPINW 298

```



```

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11585
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11585

```

```

Query Match          37.0%; Score 594; DB 15; Length 217;
Best Local Similarity 53.0%; Pred. No. 3.5e-51;
Matches 115; Conservative 38; Mismatches 62; Indels 2; Gaps 2;

```

```

QY      84  ETWRRELAAREKPYFKOLMSF-VADERSHTVYPPADQVYSWTEMCDIODYKVVILGOD 142
      1  DSKHAYLSGFEASPIYQWKEFLAKETGKRIFPKGAEYFRALDLPDEYKVVILGOD 60
      143  PYHGPNOAHGLCFSVQKVPVPPPSLVNIYKELCTDIDGFKHPGHDLSGMAKQGVLLINA 202
      61  PYHGLQAHGLCFSVQPGVRIPPSLVNIYKELQSDL-GIRPVKHGFLSEMAKQGVLLINS 119
      QY      203  VLTVEAHQANSHKDRGMEFTTDAVIKWLNVNREGVFFLLMGSYAHKKGATIDRKHHVLIQ 262
      DB      120  VLTVEARAAASHQGGMEKFTDAVIRAVNDECDHVFFLLMGSYAQKKAFFVDQRKHLVLR 179
      QY      263  AVHPSPLSAHNGFLGCKHFSKANGLLKLSGTEPINMR 299
      DB      180  SPHPSPLSAHNGFFGNGHFSKANAFLVSHGRDPIDWQ 216

```

```

RESULT 15
US-10-369-493-14855
; Sequence 14855, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14855
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14855

```

```

Query Match          37.0%; Score 594; DB 15; Length 220;
Best Local Similarity 53.0%; Pred. No. 3.6e-51;
Matches 115; Conservative 38; Mismatches 62; Indels 2; Gaps 2;

```

```

QY      84  ETWRRELAAREKPYFKOLMSF-VADERSHTVYPPADQVYSWTEMCDIODYKVVILGOD 142
      5  DSKHAYLSGFEASPIYQWKEFLAKETGKRIFPKGAEYFRALDLPDEYKVVILGOD 64

```

```

QY      143  PYHGPNOAHGLCFSVQKVPVPPPSLVNIYKELCTDIDGFKHPGHDLSGMAKQGVLLINA 202
      DB      65  PYHGLQAHGLCFSVQPGVRIPPSLVNIYKELQSDL-GIRPVKHGFLSEMAKQGVLLINS 123
      QY      203  VLTVEAHQANSHKDRGMEFTTDAVIKWLNVNREGVFFLLMGSYAHKKGATIDRKHHVLIQ 262
      DB      124  VLTVEARAAASHQGGMEKFTDAVIRAVNDECDHVFFLLMGSYAQKKAFFVDQRKHLVLR 183
      QY      263  AVHPSPLSAHNGFLGCKHFSKANGLLKLSGTEPINMR 299
      DB      184  SPHPSPLSAHNGFFGNGHFSKANAFLVSHGRDPIDWQ 220

```

```

Search completed: March 18, 2004, 00:11:37
Job time : 36 secs

```